

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/980,559C
Source: 1FW/6
Date Processed by STIC: 3/23/06

ENTERED

CRF Errors Edited by the STIC Systems Branch

Serial Number: 09/980,559C

CRF Edit Date: 3/24/06
Edited by: in

___ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

___ Corrected the SEQ ID NO. Sequence numbers edited were:

___ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

___ Deleted: ___ invalid beginning/end-of-file text ; ___ page numbers

___ Inserted mandatory headings/numeric identifiers, specifically:

___ Moved responses to same line as heading/numeric identifier, specifically:

___ Other: inserted hard returns throughout the sequence listing



IFW16

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:44:23

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

```

4 <110> APPLICANT: Yanagi, Masayuki
5     Emerson, Suzanne
6     Bukh, Jens
7     Purcell, Robert
9 <120> TITLE OF INVENTION: CLONED GENOME OF INFECTIOUS HEPATITIS C
10    VIRUSES OF GENOTYPE 2a AND USES THEREOF
13 <130> FILE REFERENCE: NIH255.001NP
15 <140> CURRENT APPLICATION NUMBER: US 09/980,559C
16 <141> CURRENT FILING DATE: 2002-05-14
18 <150> PRIOR APPLICATION NUMBER: PCT/US00/15446
19 <151> PRIOR FILING DATE: 2000-06-02
21 <150> PRIOR APPLICATION NUMBER: US 60/137,693
22 <151> PRIOR FILING DATE: 1999-06-04
24 <160> NUMBER OF SEQ ID NOS: 70
26 <170> SOFTWARE: FastSEQ for Windows Version 4.0
28 <210> SEQ ID NO: 1
29 <211> LENGTH: 9711
30 <212> TYPE: DNA
31 <213> ORGANISM: Hepatitis C virus
33 <400> SEQUENCE: 1
34 acccgccctt aataggggag acactccgcc atgaatcact cccctgtgag gaactactgt 60
35 cttcacgcag aaagcgtcta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc 120
36 cccctcccg ggagagccat agtggctcgc ggaaccggtg agtacaccgg aattgccggg 180
37 aagactgggt cctttcttgg ataaaccac tctatgcccg gccatttggg cgtgcccccg 240
38 caagactgct agccgagtag cgttgggttg cgaaaggcct tgtggtactg cctgataggg 300
39 tgcttgcgag tgccccgga ggtctcgtag accgtgcacc atgagcacia atcctaaacc 360
40 tcaaagaaaa accaaaagaa acaccaaccg tcgcccacaa gacgttaagt ttccgggcgg 420
41 cggccagatc gttggcggag tatacttggt gccgcgcagg ggccccagg tgggtgtgcg 480
42 cgcgacaagg aagacttcgg agcggtccca gccacgtgga aggcgccagc ccatccctaa 540
43 agatcggcgc tccactggca aatcctgggg aaaaccagga taccctggc ccctatacgg 600
44 gaatgaggga ctcggtggg caggatggct cctgtcccc cgaggttccc gtccctcttg 660
45 gggcccaat gacccccggc ataggtcgcg caacgtgggt aaggatcatg ataccctaac 720
46 gtgcggcttt gccgacctca tggggtacat ccctgtcggt ggcgccccgc tcggcggcgt 780
47 cgccagagct ctgcgcgatg gcgtgagagt cctggaggac ggggttaatt ttgcaacagg 840
48 gaacttaccc ggttgctcct ttctatctt ctgtgtggcc ctgtgtcct gcatcaccac 900
49 cccggtctcc gctgccgaag tgaagaacat cagtaccggc tacatggtga ctaacgactg 960
50 caccaatgac agcattacct ggcagctcca ggctgtgtgc ctccacgtcc ccgggtgcgt 1020
51 ccggtgcgag aaagtgggga atgcatctca gtgtggata ccggtctcac cgaatgtggc 1080
52 cgtgcagcgg cccggcgccc tcacgcaggg cttgcggacg cacatcgaca tggttgtgat 1140
53 gtccgccaag ctctgctctg cctctacgt gggggacctc tgcggtgggg tgatgctcgc 1200
54 agcccaaatg ttcattgtct cgccgcagca cactggttt gtccaagact gcaattgctc 1260
55 catctaccct ggtaccatca ctggacaccg catggcatgg gacatgatga tgaactggtc 1320
56 gccacaggct accatgatct tggcgtacgc gatgcgtgtc cccgagggtc ttatagacat 1380

```

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:44:23

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

```

57 cattagcggg gctcattggg gcgtcatggt cggcttgggc tacttctcta tgcagggagc 1440
58 gtgggcgaaa gtcgttggtca tccttctggt ggccgcgggg gtggacgcgc gcacccatac 1500
59 tgttgggggt tctgccgcgc agaccaccgg gcgcctcacc agcttatttg acatggggcc 1560
60 caggcagaaa atccagctcg ttaacaccaa tggcagctgg cacatcaacc gcaccgccct 1620
61 gaactgcaat gactccttgc acaccggctt tatcgcgctc ctgttctaca cccacagctt 1680
62 caactcgta ggatgtcccg aacgcagtgc cgctgccgc agtatcgagg ccttcggggt 1740
63 gggatggggc gccttgcaat atgaggataa tgtcaccaat ccagaggata tgagacccta 1800
64 ttgctggcac taccaccaa ggcagtgtgg cgtggtctcc gcgaagactg tgtgtggccc 1860
65 agtgtactgt ttcaccccca gccagtggt agtgggcacg accgacaggc ttggagcgcc 1920
66 cacttacacg tggggggaga atgagacaga tgtcttcta ttgaacagca ctcgaccacc 1980
67 gctgggggtca tgggttcggct gcacgtggat gaactcttct ggctacacca agacttgccg 2040
68 cgcaccaccc tgcctgacta gagctgactt caacgccagc acggacctgt tgtgccccac 2100
69 ggactgtttt aggaagcatc ctgataccac ttacctcaaa tgcggctctg ggccctggct 2160
70 cagccaagg tgcctgatcg actacccta caggctctgg cattaccct gcacagttaa 2220
71 ctataccatc ttcaaaataa ggatgtatgt gggaggggtt gagcacaggc tcacggctgc 2280
72 atgcaatttc actcgtgggg atcgttgcaa cttggaggac agagacagaa gtcaactgtc 2340
73 tcctttgttg cactccacca cggaatgggc cattttacct tgctcttact cggacctgcc 2400
74 cgccttgctg actggtcttc tccacctcca ccaaaacatc gtggacgtac aattcatgta 2460
75 tggcctatca cctgccctca caaaatacat cgtccgatgg gagtgggtaa tactcttatt 2520
76 cctgctctta gcggacgcca gggtttgccg ctgcttatgg atgctcatct tgttgggcca 2580
77 ggccaagca gcactagaga agctggctat cttgcacgct gcgagcgag ctagctgcaa 2640
78 tggcttcta tattttgtca tcttttctgt ggctgcttgg tacatcaagg gtcgggtagt 2700
79 ccccttagct acctattccc tcactggcct gtggtccttt agcctactgc tcctagcatt 2760
80 gcccacacag gcttatgctt atgacgcata tgtgcatggc cagataggag cggctctgct 2820
81 ggtaatgatc actctcttta ctctacccc cgggtataag acccttctca gccggttttt 2880
82 gtggtggttg tgctatcttc tgacctggg ggaagctatg gtccaggagt gggcaccacc 2940
83 tatgcagggt gcggtggcc gtgatggcat catatgggcc gtcccatat tctaccagg 3000
84 tgtggtggtt gacataacca agtggctctt ggcggtgctt gggcctgctt acctcctaaa 3060
85 aggtgctttg acgcgcgtgc cgtacttcgt cagggtcac gctctactga ggatgtgcac 3120
86 catggcaagg catctcgcg ggggcaggta cgtccagatg gcgctactag cccttggcag 3180
87 gtggactggc acttacatct atgaccacct caccctatg tcggattggg ctgctagtgg 3240
88 cctgcgggac ctggcggtcg ccgttgagcc tatcatcttc agtccgatgg agaagaaagt 3300
89 cattgtctgg ggagcggaga cagctgcttg tggggacatt ttacacggac ttcccgtgtc 3360
90 cgcccgactt ggtcgggagg tctccttgg ccagctgat ggctatacct ccaaggggtg 3420
91 gagtcttctc gccccatca ctgcttacgc ccagacagca cgtggccttt tgggcaccat 3480
92 agtggtgagc atgacggggc gcgacaagac agaacaggct ggggaaattc aggtcctgtc 3540
93 cacagtcact cagtccttcc tcggaacatc catctcgggg gttttgtgga ctgtctacca 3600
94 tggagctggc aacaagactc tggccggctc acggggtccg gtcacgcaga tgtactccag 3660
95 tgcaggggg gacttagtag ggtggcccag ccccccggg actaaatctt tggagccgtg 3720
96 cagtggtgga gcggtcgacc tgtacctggt cagcgggaac gctgatgtca tcccggctcg 3780
97 aagacgcggg gacaaacggg gacgctact ctccccgaga cctctttcca cctgaagggt 3840
98 gtccacagga gggccgggtg tatgccccag gggccacgct gtcggagtct tccgggcagc 3900
99 tgtgtgctct cggggcggtg ctaagtccat agatttcata cccgttgaga cactcgacat 3960
100 cgtcacgcgg tccccacct ttagtgacaa cagcacacca cctgctgtgc ccagacctta 4020
101 tcaggtcggg tacttgcatg ccccgactgg cagtggaaag agcaccaaaag ttccgtgtgc 4080
102 atatgctgct caggggtata aagtgctagt gcttaatccc tcagtggctg ccaccctggg 4140
103 gtttggggcg tacttgtcta aggcacatgg catcaatccc aacattagga ctggagtcag 4200
104 gactgtgacg accggggcgc ccatcacgta ctccacatat ggcaaattcc tcgccgatgg 4260
105 gggctgtgcg ggcggcgct acgacatcat catatgtgat gaatgccatg ccgtggactc 4320

```

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:44:23

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

```

106 taccaccatc cttggcatcg gaacagtcct tgatcaagca gagacagctg gggtcagact 4380
107 aactgtgctg gctacagcta cgccccctgg gtcagtgaca accccccacc ccaacataga 4440
108 ggaggtggcc cttgggcagg agggcgagat ccccttctat gggagggcga ttccctgtc 4500
109 ttacatcaag ggaggaagac atctgatctt ctgccattca aagaaaaagt gtgacgagct 4560
110 cgcggcggcc cttcggggta tgggcttgaa ctgagtgga tactacagag ggttggaagt 4620
111 ctccgtaata ccaactcagg gagacgtagt ggtcgtcgcc accgacgccc tcatgacagg 4680
112 gtatactggg gactttgact ccgtgatcga ctgcaacgta gcggtcactc aagttgtaga 4740
113 cttcagttta gaccccatat tcaccataac cacacagatt gtccctcaag acgctgtctc 4800
114 acgtagccag cgccgggggc gcacgggtag gggaagactg ggcatttata ggtatgtttc 4860
115 cactggtgag cgagcctcag gaatgtttga cagtgtagtg ctctgtgagt gctacgacgc 4920
116 agggggccga tggtagagc tcacaccatc ggagaccacc gtcaggctca gggcgatatt 4980
117 caacacgccc ggtttgcctg tgtgccaaaga ccatcttgag ttttgggagg cagttttcac 5040
118 cggcctcaca cacatagatg cccacttcct ttcccaaaca aagcaatcgg gggaaaattt 5100
119 cgcatactta acagcctacc aggtacagt gtgcgctagg gccaaagccc ccccccgctc 5160
120 ctgggacgtc atgtggaagt gtttgactcg actcaagccc aactcgtgg gcccacacc 5220
121 tctcctgtac cgcttgggct ctgttaccac cgaggtcacc ctacacatc ccgtgacgaa 5280
122 atacatcgcc acctgcatgc aagccgacct tgaggtcatg accgacatc gggcttggc 5340
123 agggggagtc ttggcgccg tcgcccgtg ttgctggcg accgggtgtg tttgcatcat 5400
124 cggccgcttg cacattaacc agcgagccgt cgttgccgcg gacaaggagg tcctctatga 5460
125 ggcttttgat gagatggagg aatgtgcctc tagggcggtc ctcatgaag aggggcagcg 5520
126 gatagccgag atgctgaagt ccaagatcca aggcctattg cagcaagctt ccaaacaagc 5580
127 tcaagacata caaccactg tgcaggcttc atggcccaag gtagaacaat tctgggcca 5640
128 acacatgtgg aacttcatta gcggcatcca atacctcga ggactatcaa cactgccagg 5700
129 gaaccctgca gtagcttcca tgatggcggt cagtgcgcgc ctaccagtc cgctgtcaac 5760
130 aagcaccact atccttctca acattttggg gggctggcta gcatcccaa ttgcaccacc 5820
131 cgcgggggccc actggcttcg ttgtcagtg cctagtggga gctgccgtag gcagtatagg 5880
132 cttaggtaag gtgctagtgg acatcctggc aggggtatgg gcgggcattt cgggggctct 5940
133 cgtcgcattc aagatcatgt ctggcgagaa gccctccatg gaggatgtcg tcaacttget 6000
134 gcctggaatt ctgtctccgg gtgccttgg agtgggagtc atctgcgcgg ccattctgcg 6060
135 ccgacacgtg ggaccggggg aaggcgccgt ccaatggatg aatagactca ttgcctttgc 6120
136 ttccagagga aatcacgtcg cccccacca ctacgtgacg gagtccgatg cgtcgcagcg 6180
137 tgtgacccaa ctacttggct ccttaccat aaccagcctg ctcaagaagc tccacaactg 6240
138 gattactgag gactgcccc tccatgcgg cggtcgtgg ctccgcgatg tgtgggactg 6300
139 ggtttgcacc atcctaacag actttaaaaa ttggctgacc tccaaattat tcccaaagat 6360
140 gcccggcctc ccttttgtct cctgtcaaaa ggggtacaag ggcgtgtggg ccggcactgg 6420
141 catcatgacc acacgggtgtc cttgcggcgc caatatctct ggcaatgtcc gcttgggctc 6480
142 catgagaatc acggggccta agacctgcat gaatatctgg caggggacct ttctatcaa 6540
143 ttgttacacg gagggccagt gcgtgcgga acccgcgcca aactttaagg tcgccatctg 6600
144 gaggggtggc gcctcagagt acgcggaggt gacgcagcac gggtcatacc actacataac 6660
145 aggactcacc actgataact tgaaagtccc ctgccaaacta cctctctccg agttcttttc 6720
146 ctgggtggac ggagtgcaga tccataggtt tgccccaca ccgaagccgt tttccggga 6780
147 tgaggtctcg ttctgcgttg ggcttaattc atttgctgct ggggtccagc ttcttgcga 6840
148 cctgaacccc gacacagacg tattgatgtc catgctaaca gatccatctc atatcacggc 6900
149 ggagactgca gcgcggcggt tagcgcgggg gtcaccccca tccgaggcaa gtcctcggc 6960
150 gagccagcta tcggcaccat cgtgcgcagc cacctgcacc acccacggca aagcctatga 7020
151 tgtggacatg gtggatgcta acctgttcat gggggcgcat gtgactcga tagagtctgg 7080
152 gtccaaagtg gtcgttcttg actctctcga ccaatggct gaagaaagga gcgacctga 7140
153 gccttcgata ccatcagaat acatgctccc caagaagagg tcccaccag ctttaccggc 7200
154 ctgggcacgg cctgattaca acccaccgct tgtggaatcg tggaaaaggc cagattacca 7260

```

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:44:23

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

```

155 accggccact gttgcgggct gtgctctccc tcctcctagg aaaacccccga cgcctccccc 7320
156 aaggaggcgc cggacagtgg gcctaagtga ggactccata ggagatgccc ttcaacagct 7380
157 ggccattaag tcctttggcc agcccccccc aagcggcgat tcaggccttt ccacgggggc 7440
158 gggcgctgcc gattccggca gtcagacgcc tcctgatgag ttggcccttt cggagacagg 7500
159 ttccatctct tccatgcccc ccctcgaggg ggagcttgga gatccagacc tggagcctga 7560
160 gcaggtagag ccccaacccc cccccaggg gggggtggca gctcccggct cggactcggg 7620
161 gtccctggtct acttgctccg aggaggacga ctccgtcgtg tgctgctcca tgtcatactc 7680
162 ctggaccggg gctctaataa ctccctgtag tccgaagag gagaagttac cgattaaccc 7740
163 cttgagcaac tcctgtttgc gatatacaca caaggtgtac tgtaccacaa caaagagcgc 7800
164 ctactaagg gctaaaaagg taacttttga taggatgcaa gtgctcgact cctactacga 7860
165 ctcagtctta aaggacatta agctagcggc ctccaaggct accgcaaggc tcctcaccat 7920
166 ggaggaggct tgccagttaa cccacaccca ttctgcaaga tctaaatatg ggtttggggc 7980
167 taaggaggct cgcagcttgt ccgggagggc cgtaaacac atcaagtccg tgtggaagga 8040
168 cctcctggag gactcagaaa caccaattcc cacaaccatt atggccaaaa atgaggtgtt 8100
169 ctgctgggac cccaccaagg ggggcaagaa agcagctcgc cttatcgttt accctgacct 8160
170 cggcgctcagg gtctgcgaga agatggccct ttatgacatt acacaaaaac ttccctcaggc 8220
171 ggtgatgggg gcttcttatg gattccagta ttccccgcgt cagcgggtag agtttctctt 8280
172 gaaagcatgg gcggaaaaga aggaccctat gggtttttcg tatgataccc gatgctttga 8340
173 ctcaaccgtc actgagagag acatcaggac tgaggagtcc atatatcggg cctgctcctt 8400
174 gcccaggagg gccacactg ccatacactc gtaactgag agactttacg tgggaggggc 8460
175 tatgttcaac agcaagggcc aaacctgcgg gtacaggcgt tgccgcgcca gcgggggtgct 8520
176 caccactagc atggggaaca ccatcacatg ctacgtgaaa gccttagcgg cttgtaaagc 8580
177 tgcagggata atcgcgccca caatgctggt atgcggcgat gacttggttg tcatctcaga 8640
178 aagccagggg accgaggagg acgagcggaa cctgagagcc ttcacggagg ctatgaccag 8700
179 gtattctgcc cctcctggtg accccccccag accggagtat gatctggagc tgataacatc 8760
180 ttgctcctca aatgtgtctg tggcgctggg cccacaaggc cgccgcagat actacctgac 8820
181 cagagaccct accactccaa tcgcccgggc tgccctggga acagttagac actcccctgt 8880
182 caattcatgg ctgggaaaca tcatccagta cgccccgacc atatgggctc gcatggctct 8940
183 gatgacacac ttcttctcca ttctcatggc tcaagacacg ctggaccaga acctcaactt 9000
184 tgagatgtac ggagcggtgt actccgtgag tcccttgga ctcccagcta taattgaaag 9060
185 gttacatggg cttgacgctt tttctctgca cacatacact cccacgaac tgacacgggt 9120
186 ggcttcagcc ctcaaaaac ttggggcgcc accctcaga gcgtggaaga gccgggcacg 9180
187 tgcagtcagg gcgtccctca tctcccgtag ggggagagcg gccgtttgcg gtcgatatct 9240
188 cttcaattgg gcggtgaaga ccaagctcaa actcactcca ttgccggaag cgcgccctct 9300
189 ggatttatcc agctggttca ccgtcggcgc cggcgggggc gacatttatc acagcgtgtc 9360
190 gcgtgcccga ccccgcttat tgctctttgg cctactccta cttttttagt gggtaggcct 9420
191 tttcctactc cccgctcggg agagcggcac acattagcta cactccatag ctaactgtcc 9480
192 cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 9540
193 tttttttttt tttttttttt tttttctttt tttctctttt ccttctttct taccttattt 9600
194 tactttcttt cctggtggct ccactcttag cctagtcacg gctagctgtg aaagggtccgt 9660
195 gagccgcatg actgcagaga gtgccgtaac tggctctctt gcagatcatg t 9711

```

197 <210> SEQ ID NO: 2

198 <211> LENGTH: 3033

199 <212> TYPE: PRT

200 <213> ORGANISM: Hepatitis C virus

202 <400> SEQUENCE: 2

203 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn

204 1 5 10 15

205 Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:44:23

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

```

206          20          25          30
207 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
208          35          40          45
209 Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
210          50          55          60
211 Ile Pro Lys Asp Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly
212 65          70          75          80
213 Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
214          85          90          95
215 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro
216          100          105          110
217 Arg His Arg Ser Arg Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys
218          115          120          125
219 Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Val Val Gly Ala Pro Leu
220          130          135          140
221 Gly Gly Val Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp
222 145          150          155          160
223 Gly Val Asn Phe Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile
224          165          170          175
225 Phe Leu Leu Ala Leu Leu Ser Cys Ile Thr Thr Pro Val Ser Ala Ala
226          180          185          190
227 Glu Val Lys Asn Ile Ser Thr Gly Tyr Met Val Thr Asn Asp Cys Thr
228          195          200          205
229 Asn Asp Ser Ile Thr Trp Gln Leu Gln Ala Ala Val Leu His Val Pro
230          210          215          220
231 Gly Cys Val Pro Cys Glu Lys Val Gly Asn Ala Ser Gln Cys Trp Ile
232 225          230          235          240
233 Pro Val Ser Pro Asn Val Ala Val Gln Arg Pro Gly Ala Leu Thr Gln
234          245          250          255
235 Gly Leu Arg Thr His Ile Asp Met Val Val Met Ser Ala Thr Leu Cys
236          260          265          270
237 Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ala
238          275          280          285
239 Gln Met Phe Ile Val Ser Pro Gln His His Trp Phe Val Gln Asp Cys
240          290          295          300
241 Asn Cys Ser Ile Tyr Pro Gly Thr Ile Thr Gly His Arg Met Ala Trp
242 305          310          315          320
243 Asp Met Met Met Asn Trp Ser Pro Thr Ala Thr Met Ile Leu Ala Tyr
244          325          330          335
245 Ala Met Arg Val Pro Glu Val Ile Ile Asp Ile Ile Ser Gly Ala His
246          340          345          350
247 Trp Gly Val Met Phe Gly Leu Ala Tyr Phe Ser Met Gln Gly Ala Trp
248          355          360          365
249 Ala Lys Val Val Val Ile Leu Leu Leu Ala Ala Gly Val Asp Ala Arg
250          370          375          380
251 Thr His Thr Val Gly Gly Ser Ala Ala Gln Thr Thr Gly Arg Leu Thr
252 385          390          395          400
253 Ser Leu Phe Asp Met Gly Pro Arg Gln Lys Ile Gln Leu Val Asn Thr
254          405          410          415

```

VERIFICATION SUMMARY

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:44:24

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

**Raw Sequence Listing before editing
(for reference only)**



IFW16

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:31:17

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

4 <110> APPLICANT: Yanagi, Masayuki
 5 Emerson, Suzanne
 6 Bukh, Jens
 7 Purcell, Robert
 9 <120> TITLE OF INVENTION: CLONED GENOME OF INFECTIOUS HEPATITIS C
 10 VIRUSES OF GENOTYPE 2a AND USES THEREOF
 13 <130> FILE REFERENCE: NIH255.001NP
 15 <140> CURRENT APPLICATION NUMBER: US 09/980,559C
 16 <141> CURRENT FILING DATE: 2002-05-14
 18 <150> PRIOR APPLICATION NUMBER: PCT/US00/15446
 19 <151> PRIOR FILING DATE: 2000-06-02
 21 <150> PRIOR APPLICATION NUMBER: US 60/137,693
 22 <151> PRIOR FILING DATE: 1999-06-04
 24 <160> NUMBER OF SEQ ID NOS: 70
 26 <170> SOFTWARE: FastSEQ for Windows Version 4.0

*see
pp 2-8*
**Does Not Comply
Corrected Diskette Needed**

ERRORED SEQUENCES

28 <210> SEQ ID NO: 1
 29 <211> LENGTH: 9711
 30 <212> TYPE: DNA
 31 <213> ORGANISM: Hepatitis C virus
 33 <400> SEQUENCE: 1

see P.2

34 accgccccct aataggggag acactccgcc atgaatcact cccctgtgag gaactactgt 60
 35 cttcacgcag aaagcgtcta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc 120
 36 cccctccccg ggagagccat agtggctctgc ggaaccggtg agtacaccgg aattgccggg 180
 37 aagactgggt cctttcttgg ataaaccac tctatgcccg gccatttggg cgtgcccccg 240
 38 caagactgct agccgagtag cgttgggttg cgaaaggcct tgtggtactg cctgataagg 300
 39 tgcttgcgag tgccccggga ggtctcgtag accgtgcacc atgagcaca atcctaaacc 360
 40 tcaaagaaaa accaaaagaa acaccaaccg tcgcccacaa gacgttaagt ttccgggcgg 420
 41 cggccagatc gttggcggag tatacttggt gccgcgcagg ggccccagg tgggtgtgcg 480
 42 cgcgacaagg aagacttcgg agcgggtccca gccacgtgga aggcgccagc ccatccctaa 540
 43 agatcggcgc tccactggca aatcctgggg aaaaccagga taccctggc ccctatacgg 600
 44 gaatgaggga ctcggtcggg caggatggct cctgtccccc cgagggtccc gtcctctctg 660
 45 gggccccaat gacccccggc ataggctcgc caacgtgggt aagggtcatc ataccctaac 720
 46 gtgcggcttt gccgacctca tggggtacat cctgtcgtg ggcgccccgc tcggcggcgt 780
 47 cgccagagct ctgcgcgatg gcgtgagagt cctggaggac ggggttaatt ttgcaacagg 840
 48 gaacttacct ggttgctcct tttctatct cttgctggcc ctgctgtcct gcataccac 900
 49 cccggtctcc gctgccgaag tgaagaacat cagtaccggc tacatggtga ctaacgactg 960
 50 caccaatgac agcattacct ggcagctcca ggetgctgtc ctccacgtcc ccgggtgctg 1020
 51 cccgtgcgag aaagtgggga atgcatctca gtgctggata ccggtctcac cgaatgtggc 1080
 52 cgtgcagcgg cccggcgccc tcacgcaggg cttgcggacg cacatcgaca tggttgtgat 1140

2

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:31:17

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

```

151 tgtggacatg gtggatgcta acctgttcat ggggggcgat gtgactcgga tagagtctgg 7080
152 gtccaaagtg gtcgttctgg actctctcga cccaatgggtc gaagaaagga gcgaccttga 7140
153 gccttcgata ccatcagaat acatgctccc caagaagagg tccccaccag ctttaccggc 7200
154 ctgggcacgg cctgattaca acccaccgct tgtggaatcg tggaaaaggc cagattacca 7260
155 accggccact gttgcgggct gtgctctccc tcttcctagg aaaaccccca cgccctcccc 7320
156 aaggaggcgc cggacagtgg gcctaagtga ggactccata ggagatgccc ttcaacagct 7380
157 ggccattaag tcttttggcc agcccccccc aagcggcgat tcaggccttt ccacgggggc 7440
158 gggcgctgcc gattccggca gtcagacgcc tctgatgag ttggcccttt cggagacagg 7500
159 ttccatctct tccatgcccc cctcagagg ggagcttgga gatccagacc tggagcctga 7560
160 gcaggtagag ccccaacccc cccccagggg gggggtggca gctcccggct cggactcggg 7620
161 gtcctgggtct acttgctccg aggaggacga ctccgtcgtg tgctgctcca tgcatactc 7680
162 ctggaccggg gctctaataa ctcttgtag tcccgagag gagaagttac cgattaaccc 7740
163 cttgagcaac tccctgttgc gatatacaca caaggtgtac tgtaccacaa caaagagcgc 7800
164 ctactaagg gctaaaaagg taacttttga taggatgcaa gtgctcgact cctactacga 7860
165 ctcagcttta aaggacatta agctagcggc ctccaaggct accgcaaggc tctcaccat 7920
166 ggaggaggct tgccagttaa ccccacccca ttctgcaaga tctaaatatg ggtttggggc 7980
167 taaggaggct cgcagcttgt ccgggagggc cgtaaacac atcaagtcg tgtggaagga 8040
168 cctcctggag gactcagaaa caccaattcc cacaaccatt atggccaaaa atgaggtgtt 8100
169 ctgctgggac cccaccaagg ggggcaagaa agcagctcgc cttatcgttt accctgacct 8160
170 cggcgctcagg gtctgcgaga agatggccct ttatgacatt acacaaaaac ttcctcaggc 8220
171 ggtgatgggg gcttcttatg gattccagta tcccccgct cagcgggtag agtttctctt 8280
172 gaaagcatgg gcggaaaaga aggaccctat ggggttttct tatgataccc gatgctttga 8340
173 ctcaaccgtc actgagagag acatcaggac tgaggagtc atatatcggg cctgctcctt 8400
174 gcccgaggag gccacactg ccatacactc gctaactgag agactttacg tgggaggggc 8460
175 tatgttcaac agcaagggcc aaacctgcgg gtacaggcgt tgccgcgcca gcggggtgct 8520
176 caccactagc atgggggaaca ccatcacatg ctacgtgaaa gccttagcgg cttgtaaagc 8580
177 tgcagggata atcgcgccca caatgctggg atgcggcgat gacttggttg tcatctcaga 8640
178 aagccagggg accgaggagg acgagcggaa cctgagagcc ttcacggagg ctatgaccag 8700
179 gtattctgcc cctcctgggtg accccccccag accggagtat gatctggagc tgataacatc 8760
180 ttgctcctca aatgtgtctg tggcgctggg ccacaaaggc cgccgcagat actacctgac 8820
181 cagagaccct accactccaa tcgcccgggc tgccctggga acagttagac actcccctgt 8880
182 caattcatgg ctgggaaaca tcatccagta cgccccgacc atatgggctc gcatggctct 8940
183 gatgacacac ttcttctcca ttctcatggc tcaagacacg ctggaccaga acctcaactt 9000
184 tgagatgtac ggagcgggtgt actccgtgag tcccttggac ctcccagcta taattgaaag 9060
185 gttacatggg cttgacgctt tttctctgca cacatacact cccacgaac tgacacgggt 9120
186 ggcttcagcc ctcaaaaaac ttggggcgcc accctcaga gcgtggaaga gccgggcacg 9180
187 tgcagtcagg gcgtccctca tctcccggtg ggggagagcg gccgtttgcg gtcgatatct 9240
188 cttcaattgg gcggtgaaga ccaagctcaa actcactcca ttgccggaag cgcgcctcct 9300
189 ggatttatcc agctgggttca ccgtcggcgc cggcgggggc gacatttatc acagcgtgtc 9360
190 gcgtgcccga ccccgcttat tgctctttgg cctactccta cttttttagg gggtaggcct 9420
191 tttcctactc cccgctcggg agagcggcac acattagcta cactccatag ctaactgtcc 9480
192 cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 9540
193 tttttttttt tttttttttt tttttctttt tttctctttt cttctctttt taccttattt 9600
194 tactttcttt cctggtgggt ccatcttagc cctagtcacg gctagctgtg aaaggctccg 9660

```

Insert
a hard
return

```

E--> 195 gagccgcatg actgcagaga gtgccgtaac tgggtctctct gcagatcatg t 9711<210> 2
196 <211> LENGTH: 3033
197 <212> TYPE: PRT
198 <213> ORGANISM: Hepatitis C virus

```

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:31:17

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

742 agtacctctt caactgggca gtaagaacaa agctcaaact cactccaata gcggccgctg 9240
743 gccggctgga cttgtccggg tgggtcacgg ctggctacag cgggggagac atttatcaca 9300
744 gcgtgtctca tgcccgcccc cgctgggtct ggttttgctt actcctgctc gctgcagggg 9360
745 taggcattcta cctcctcccc aaccgatgaa ggttggggta aacactccgg cctcttaagc 9420
746 catttcctgt tttttttttt tttttttttt tttttttctt tttttttttc tttcctttcc 9480
747 ttcttttttt cctttctttt tcccttcttt aatggtggct ccatcttagc cctagtcacg 9540
748 gctagctgtg aaagggtccg gagccgcatg actgcagaga gtgctgatac tggcctctct 9600

E--> 749 gcagatcatg t

9611<210> 4

750 <211> LENGTH: 3015

751 <212> TYPE: PRT

752 <213> ORGANISM: Hepatitis C virus

E--> 754 <400> SEQUENCE: 4

755 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
756 1 5 10 15
757 Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
758 20 25 30
759 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
760 35 40 45
761 Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
762 50 55 60
763 Ile Pro Lys Asp Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly
764 65 70 75 80
765 Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
766 85 90 95
767 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro
768 100 105 110
769 Arg His Arg Ser Arg Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys
770 115 120 125
771 Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Val Val Gly Ala Pro Leu
772 130 135 140
773 Gly Gly Val Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp
774 145 150 155 160
775 Gly Val Asn Phe Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile
776 165 170 175
777 Phe Leu Leu Ala Leu Leu Ser Cys Ile Thr Thr Pro Val Ser Ala Ala
778 180 185 190
779 Glu Val Lys Asn Ile Ser Thr Gly Tyr Met Val Thr Asn Asp Cys Thr
780 195 200 205
781 Asn Asp Ser Ile Thr Trp Gln Leu Gln Ala Ala Val Leu His Val Pro
782 210 215 220
783 Gly Cys Val Pro Cys Glu Lys Val Gly Asn Ala Ser Gln Cys Trp Ile
784 225 230 235 240
785 Pro Val Ser Pro Asn Val Ala Val Gln Arg Pro Gly Ala Leu Thr Gln
786 245 250 255
787 Gly Leu Arg Thr His Ile Asp Met Val Val Met Ser Ala Thr Leu Cys
788 260 265 270
789 Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ala
790 275 280 285

insert
hard
return

4 - - - -

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:31:17

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

```

1284 tctgtgaaag tgcggggggc caggaggacg cggcgagcct gagagccttc acggaggcta 8640
1285 tgaccaggta ctccgcccc cccggggacc cccacaacc agaatacgac ttggagctta 8700
1286 taacatcatg ctccccaac gtgtcagtcg ccacagcagg cgctggaaag aggggtctact 8760
1287 accttaccog tgacctaca accccctcg cgagagccgc gtgggagaca gcaagacaca 8820
1288 ctccagtcaa ttctggcta ggcaacataa tcatgtttgc cccacactg tgggcgagga 8880
1289 tgatactgat gacctattc tttagcgtcc tcatagccag ggatcagctt gaacaggctc 8940
1290 ttaactgtga gatctacgga gctgtctact ccatagaacc actggatcta cctccaatca 9000
1291 ttcaaagact ccatggcctc agcgcatttt cactccacag ttactctcca ggtgaaatca 9060
1292 atagggtggc cgcagcctc agaaaacttg ggggtccgcc cttgcgagct tggagacacc 9120
1293 gggcccgagg cgtccgcgct aggtctctgt ccagaggagg cagggctgct atatgtggca 9180
1294 agtacctctt caactgggca gtaagaacaa agctcaaact cactccaata gcggccgctg 9240
1295 gccggctgga cttgtccggt tgggtcacgg ctggctacag cgggggagac atttatcaca 9300
1296 gcgtgtctca tgcccgccc cgtgtgttct ggttttgctt actcctgctc gctgcagggg 9360
1297 taggcatcta cctcctccc aaccgatgaa ggttggggta aacactccgg cctcttaagc 9420
1298 catttctgtt ttttttttt ttttttttt ttttttttt ttttttttt tttcctttcc 9480
1299 ttcttttttt cctttctttt tcccttcttt aatggtggct ccatcttagc cctagtcacg 9540
1300 gctagctgtg aaaggctcgt gagccgcatg actgcagaga gtgctgatac tggcctctct 9600

```

hand
return

E--> 1301 gcagatcatg t 9611<210> 6
 1302 <211> LENGTH: 3015
 1303 <212> TYPE: PRT
 1304 <213> ORGANISM: Hepatitis C virus

E--> 1306 <400> SEQUENCE: 6

1307	Met	Ser	Thr	Asn	Pro	Lys	Pro	Gln	Arg	Lys	Thr	Lys	Arg	Asn	Thr	Asn
1308	1				5					10					15	
1309	Arg	Arg	Pro	Gln	Asp	Val	Lys	Phe	Pro	Gly	Gly	Gly	Gln	Ile	Val	Gly
1310				20					25					30		
1311	Gly	Val	Tyr	Leu	Leu	Pro	Arg	Arg	Gly	Pro	Arg	Leu	Gly	Val	Arg	Ala
1312			35					40					45			
1313	Thr	Arg	Lys	Thr	Ser	Glu	Arg	Ser	Gln	Pro	Arg	Gly	Arg	Arg	Gln	Pro
1314		50					55					60				
1315	Ile	Pro	Lys	Asp	Arg	Arg	Ser	Thr	Gly	Lys	Ser	Trp	Gly	Lys	Pro	Gly
1316	65					70				75					80	
1317	Tyr	Pro	Trp	Pro	Leu	Tyr	Gly	Asn	Glu	Gly	Leu	Gly	Trp	Ala	Gly	Trp
1318				85						90					95	
1319	Leu	Leu	Ser	Pro	Arg	Gly	Ser	Arg	Pro	Ser	Trp	Gly	Pro	Asn	Asp	Pro
1320			100						105					110		
1321	Arg	His	Arg	Ser	Arg	Asn	Val	Gly	Lys	Val	Ile	Asp	Thr	Leu	Thr	Cys
1322			115					120					125			
1323	Gly	Phe	Ala	Asp	Leu	Met	Gly	Tyr	Ile	Pro	Val	Val	Gly	Ala	Pro	Leu
1324		130					135					140				
1325	Gly	Gly	Val	Ala	Arg	Ala	Leu	Ala	His	Gly	Val	Arg	Val	Leu	Glu	Asp
1326	145					150				155					160	
1327	Gly	Val	Asn	Phe	Ala	Thr	Gly	Asn	Leu	Pro	Gly	Cys	Ser	Phe	Ser	Ile
1328				165						170					175	
1329	Phe	Leu	Leu	Ala	Leu	Leu	Ser	Cys	Ile	Thr	Thr	Pro	Val	Ser	Ala	Ala
1330			180						185					190		
1331	Glu	Val	Lys	Asn	Ile	Ser	Thr	Gly	Tyr	Met	Val	Thr	Asn	Asp	Cys	Thr
1332			195					200					205			

5

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:31:18

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

```

1826 ggaaagacct tctggaagac agtghtaacac caatagacac taccatcatg gccaaagaacg 8040
1827 aggtttttctg cgttcagcct gagaagggggg gtcgtaagcc agctcgtctc atcgtgttcc 8100
1828 ccgacctggg cgtgcgcgtg tgcgagaaga tggccctgta cgacgtgggt agcaagctcc 8160
1829 ccctggccgt gatgggaagc tcctacggat tccaatactc accaggacag cgggttgaat 8220
1830 tcctcgtgca agcgtggaag tccaagaaga ccccgatggg gttctcgtat gatacccgct 8280
1831 gttttgactc cacagtcact gagagcgaca tccgtacgga ggaggcaatt taccaatgtt 8340
1832 gtgacctgga cccccaagcc cgcgtggcca tcaagtccct cactgagagg ctttatgttg 8400
1833 ggggccctct taccaattca aggggggaaa actgcggcta ccgcagggtgc cgcgcgagcg 8460
1834 gcgtactgac aactagctgt ggtaacaccc tcaattgcta catcaaggcc cgggcagcct 8520
1835 gtcgagccgc agggctccag gactgcacca tgctcgtgtg tggcgacgac ttagtcgtta 8580
1836 tctgtgaaag tgcgggggtc caggaggacg cggcgagcct gagagccttc acggaggcta 8640
1837 tgaccaggta ctccgcccc cccggggacc ccccaaac agaatagcag ttggagctta 8700
1838 taacatcatg ctctccaac gtgtcagtcg cccacgacgg cgctggaaag agggctact 8760
1839 accttaccg tgacctaca cccccctcg cgagagccgc gtgggagaca gcaagacaca 8820
1840 ctccagtcaa ttcttggtta ggcaacataa tcatgtttgc cccacactg tggcgaggga 8880
1841 tgatactgat gacctttt tttagcgtcc tcatagccag ggatcagctt gaacaggctc 8940
1842 ttaactgtga gatctacgga gcctgctact ccatagaacc actggatcta cctccaatca 9000
1843 ttcaaagact ccatggcctc agcgcatttt cactccacag ttactctcca ggtgaaatca 9060
1844 atagggtggc cgcattgctc agaaaacttg ggggtcccgcc cttgcgagct tggagacacc 9120
1845 gggcccgagg cgtccgcgct aggtctctgt ccagaggagg cagggtctgt atatgtggca 9180
1846 agtacctctt caactgggca gtaagaacaa agctcaaact cactccaata gcggccgctg 9240
1847 gccggtgga cttgtccggt tggttcacgg ctggctacag cgggggagac atttatcaca 9300
1848 gcgtgtctca tgcccgccc cgtggttct ggttttgct actcctgctc gctgcagggg 9360
1849 taggcatcta cctctcccc aaccgatgaa ggttggggta aacactccgg cctcttaagc 9420
1850 catttctgt ttttttttt ttttttttt ttttttttt ttttttttt tttctttcc 9480
1851 ttcttttttt cttttcttt tcccttcttt aatggtggct ccatcttagc cctagtcacg 9540
1852 gctagctgtg aaaggctcgt gagccgcatg actgcagaga gtgctgatac tggcctctct 9600

```

hard
return

E--> 1853 gcagatcatg t 9611<210> 8

1854 <211> LENGTH: 3015

1855 <212> TYPE: PRT

1856 <213> ORGANISM: Hepatitis C virus

E--> 1858 <400> SEQUENCE: 8

```

1859 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
1860 1 5 10 15
1861 Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
1862 20 25 30
1863 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
1864 35 40 45
1865 Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
1866 50 55 60
1867 Ile Pro Lys Asp Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly
1868 65 70 75 80
1869 Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
1870 85 90 95
1871 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro
1872 100 105 110
1873 Arg His Arg Ser Arg Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys
1874 115 120 125

```

6

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:31:18

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

```

2368 ccttggccga gcttggccacc aaaagttttg gcagctcctc aacttccggc attacgggcg 7440
2369 acaatacgac aacatcctct gagcccgccc cttctggctg cccccccgac tccgacgttg 7500
2370 agtccatttc ttccatgccc cccctggagg gggagcctgg ggatccggat ctacgacgac 7560
2371 ggtcatggtc gacggtcagt agtggggccg acacggaaga tgtcgtgtgc tgcataatgt 7620
2372 cttattcctg gacaggcgca ctgcgcaccc cgtgcgctgc ggaagaacaa aaactgccc 7680
2373 tcaacgcact gagcaactcg ttgctacgcc atcacaatct ggtgtattcc accacttcac 7740
2374 gcagtgtctg ccaaaggcag aagaaagtca catttgacag actgcaagtt ctggacagcc 7800
2375 attaccagga cgtgctcaag gaggtcaaag cagcggcgctc aaaagtgaag gctaacttgc 7860
2376 tatccgtaga ggaagcttgc agcctgacgc cccacattc agccaaatcc aagtttggct 7920
2377 atggggcaaa agacgtccgt tgccatgcca gaaaggccgt agcccacatc aactccgtgt 7980
2378 ggaaagacct tctggaagac agtghtaacac caatagacac taccatcatg gccaagaacg 8040
2379 aggttttctg cgttcagcct gagaaggggg gtcgtaagcc agctcgtctc atcgtgttcc 8100
2380 ccgacctggg cgtgcgcgtg tgcgagaaga tggccctgta cgacgtggtt agcaagctcc 8160
2381 ccctggccgt gatgggaagc tcctacggat tccaatactc accaggacag cgggttgaat 8220
2382 tcctcgtgca agcgtggaag tccaagaaga ccccgatggg gttctcgtat gatacccgct 8280
2383 gttttgactc cacagtcact gagagcgaca tccgtacgga ggaggcaatt taccaatgtt 8340
2384 gtgacctgga cccccaagcc cgcgtggcca tcaagtccct cactgagagg ctttatgttg 8400
2385 ggggccctct taccaattca aggggggaaa actgcggcta ccgcaggtgc cgcgcgagcg 8460
2386 gcgtactgac aactagctgt ggtaacaccc tcacttgcta catcaaggcc cgggcagcct 8520
2387 gtcgagccgc agggctccag gactgcacca tgctcgtgtg tggcgacgac ttagtcgtta 8580
2388 tctgtgaaaag tgcgggggtc caggaggacg cggcgagcct gagagccttc acggaggcta 8640
2389 tgaccaggta ctccgcccc cccggggacc cccacaacc agaatacgac ttggagctta 8700
2390 taacatcatg ctctccaac gtgtcagtcg cccacgacgg cgctggaaag agggctact 8760
2391 accttaccg tgacctaca acccccctcg cgagagccgc gtgggagaca gcaagacaca 8820
2392 ctccagtcaa ttctggcta ggcaacataa tcatgtttgc cccacactg tgggcgagga 8880
2393 tgatactgat gacccatttc tttagcgtcc tcatagccag ggatcagctt gaacaggctc 8940
2394 ttaactgtga gatctacgga gcctgctact ccatagaacc actggatcta cctccaatca 9000
2395 ttcaaagact ccatggcctc agcgcatttt cactccacag ttactctcca ggtgaaatca 9060
2396 atagggtggc cgcagcctc agaaaacttg gggctccgcc cttgcgagct tggagacacc 9120
2397 gggcccgag cgtccgcgt aggtctctgt ccagaggagg cagggtctgt atatgtggca 9180
2398 agtacctctt caactgggca gtaagaacaa agctcaaact cactccaata gcggccgctg 9240
2399 gccggctgga cttgtccggt tgggttcacgg ctggctacag cgggggagac atttatcaca 9300
2400 gcgtgtctca tgcccggccc cgtggttct ggttttgcct actcctgctc gctgcagggg 9360
2401 taggcatcta cctcctcccc aaccgatgaa gggtggggta aacactccgg cctcttaagc 9420
2402 catttctctg tttttttttt tttttttttt tttttttctt tttttttttc tttcctttcc 9480
2403 ttcttttttt cttttctttt tcccttcttt aatggtggct ccatcttagc cctagtacg 9540
2404 gctagctgtg aaaggctcgt gagccgcatg actgcagaga gtgctgatac tggcctctct 9600

```

E--> 2405 gcagatcatg t 9611<210> 10

2406 <211> LENGTH: 3015

2407 <212> TYPE: PRT

2408 <213> ORGANISM: Hepatitis C virus

E--> 2410 <210> SEQ ID NO:

E--> 2410 <400> SEQUENCE: 10

```

2411 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
2412 1 5 10 15
2413 Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
2414 20 25 30
2415 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
2416 35 40 45

```

7

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:31:18

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

```

3632 ggggtgcgcct acacaggttt gcgccccctt gcaagccctt gctgcgggag gaggtatcat 6780
3633 tcagagtagg actccacgag taccgggtgg ggtcgcaatt accttgcgag cccgaaccgg 6840
3634 acgtagccgt gttgacgtcc atgctcactg atccctccca tataacagca gaggcggccg 6900
3635 ggagaagggtt ggcgagaggg tcacccccctt ctatggccag ctccctcggt agccagctgt 6960
3636 ccgctccatc tctcaaggca acttgaccg ccaaccatga ctccctgac gccgagctca 7020
3637 tagaggctaa cctcctgtgg aggcaggaga tgggcggcaa catcaccagg gttgagtcag 7080
3638 agaacaaagt ggtgattctg gactccttcg atccgcttgt ggcagaggag gatgagcggg 7140
3639 aggtctccgt acctgcagaa attctgcgga agtctcgag attcgcccg gccctgcccg 7200
3640 tctgggcgcg gccggactac aacccccgc tagtagagac gtggaaaaag cctgactacg 7260
3641 aaccacctgt ggtccatggc tgcccgtac caccctccag gtccctcct gtgcctccgc 7320
3642 ctcggaaaaa gcgtacggtg gtccctaccg aatcaaccct atctactgcc ttggccgagc 7380
3643 ttgccaccaa aagttttggc agctcctcaa ctccggcat tacgggcgac aatacgacaa 7440
3644 catcctctga gcccgcctt tctggctgcc ccccgactc cgacgttgag tcctattctt 7500
3645 ccatgcccc cctggagggg gagcctgggg atccggatct cagcgacggg tcatggtcga 7560
3646 cggtcagtag tggggccgac acggaagatg tcgtgtgtct ctcaatgtct tattcctgga 7620
3647 caggcgact cgtcaccctg tgcgtgcgg aagaacaaaa actgcccac aacgcactga 7680
3648 gcaactcgtt gctacgccat cacaatctgg tgtattccac cacttcacgc agtgcttgcc 7740
3649 aaaggcagaa gaaagtcaca ttgacagac tgcaagttct ggacagccat taccaggacg 7800
3650 tgctcaaggga ggtcaaagca gcggcggtcaa aagtgaaggc taacttgcta tccgtagagg 7860
3651 aagcttgtag cctgacgccc ccacattcag ccaaatccaa gtttggtat ggggcaaaa 7920
3652 acgtccgttg ccatgccaga aaggccgtag cccacatcaa ctccgtgtgg aaagacctt 7980
3653 tggagacag tgtaacacca atagacacta ccatcatggc caagaacgag gttttctgcg 8040
3654 ttcagcctga gaaggggggt cgtaagccag ctctgtctcat cgtgttcccc gacctgggag 8100
3655 tgcgcgtgtg cgagaagatg gccctgtacg acgtggttag caagctcccc ctggccgtga 8160
3656 tgggaagctc ctacggattc caatactcac caggacagcg ggttgaattc ctctgcaag 8220
3657 cgtggaagtc caagaagacc ccgatgggg tctcgtatga taccgctgt tttgactcca 8280
3658 cagtcactga gagcgacatc cgtacggagg aggcaattta ccaatgttgt gacctggacc 8340
3659 cccaagcccc cgtggccatc aagtccctca ctgagaggct ttatgttggg ggccctctta 8400
3660 ccaattcaag gggggaaaac tgccgctacc gcagggtgcc cgcgagcggc gtactgacaa 8460
3661 ctagctgtgg taacaccctc acttgctaca tcaaggcccg ggcagcctgt cgagccgcag 8520
3662 ggctccagga ctgcaccatg ctctgtgtgt gcgacgactt agtcgttatc tgtgaaagt 8580
3663 cgggggtcca ggaggacgcg gcgagcctga gaccttcac ggaggctatg accaggact 8640
3664 ccgccccccc cggggacccc ccacaaccag aatacgactt ggagcttata acatcatgct 8700
3665 cctccaacgt gtcagtcgcc cagcagggcg ctggaaagag ggtctactac cttaccctgt 8760
3666 accctacaac cccctcgcg agagccgctt gggagacagc aagacacact ccagtcaatt 8820
3667 cctggctagg caacataatc atgtttgccc ccacactgtg ggcgaggatg atactgatga 8880
3668 cccatttctt tagcgtctc atagccaggg atcagcttga acaggctctt aactgtgaga 8940
3669 tctacggagc ctgctactcc atagaaccac tggatctacc tccaatcatt caaagactcc 9000
3670 atggcctcag cgcattttca ctccacagt actctccagg tgaaatcaat aggggtggcg 9060
3671 catgcctcag aaaacttggg gtcccgcct tgcgagcttg gagacaccgg gcccgagcg 9120
3672 tccgcgtag gcttctgtcc agaggaggca gggctgccat atgtggcaag tacctcttca 9180
3673 actgggcagt aagaacaaag ctcaaactca ctccaatagc ggccgctggc cggctggact 9240
3674 tgtccggttg gttcacggct ggctacagcg ggggagacat ttatcacagc gtgtctcatg 9300
3675 cccggccccg ctggttctgg ttttgctac tctgtctgc tgccgggta ggcattacc 9360
3676 tcctcccaaa ccgatgaagg ttggggtaaa cactccggcc tcttaagcca tttcctgttt 9420
3677 tttttttttt tttttttttt tttttcttt ttttttctt tcctttcctt ctttttttcc 9480
3678 tttctttttt cttcttttaa tgggtggctcc atcttagccc tagtcacggc tagctgtgaa 9540
3679 aggtccgtga gccgcatgac tgcagagagt gctgatactg gcctctctgc agatcatgt 9599<210> 68
3680 <211> LENGTH: 3011

```

E-->

8

RAW SEQUENCE LISTING

DATE: 03/24/2006

PATENT APPLICATION: US/09/980,559C

TIME: 18:31:18

Input Set : A:\PTO.txt

Output Set: N:\CRF4\03242006\I980559C.raw

```
4223 cgcgctaagc tactgtccca ggggggggagg gccgccactt gtggcagata cctctttaac 9180
4224 tgggcagtaa ggaccaagct taaactcact ccaatcccgg ccgcgtccca gctggacttg 9240
4225 tctggctggg tcgtcgctgg ttacagcggg ggagacatat atcacagcct gtctcgtgcc 9300
4226 cgaccccgtt gggtttccgtt gtgcctactc ctactttctg taggggtagg catttacctg 9360
4227 ctccccaacc gatgaacggg gagctaacca ctccaggcct taagccattt cctgtttttt 9420
4228 tttttttttt tttttttttt ttcttttttt ttttctttcc tttcctttct tttttccttt 9480
4229 ctttttccct tctttaatgg tggctccatc ttagccctag tcacggctag ctgtgaaagg 9540 ✓
E--> 4230 tccgtgagcc gcatgactgc agagagtgtc gatactggcc tctctgcaga tcatgt 9596<210> 70
4231 <211> LENGTH: 3010
4232 <212> TYPE: PRT
4233 <213> ORGANISM: Hepatitis C virus
E--> 4235 <210> SEQ ID NO:
E--> 4235 <400> SEQUENCE: 70
4236 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
4237 1 5 10 15
4238 Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
4239 20 25 30
4240 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
4241 35 40 45
4242 Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
4243 50 55 60
4244 Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Ala Trp Ala Gln Pro Gly
4245 65 70 75 80
4246 Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
4247 85 90 95
4248 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro
4249 100 105 110
4250 Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys
4251 115 120 125
4252 Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu
4253 130 135 140
4254 Gly Gly Ala Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp
4255 145 150 155 160
4256 Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile
4257 165 170 175
4258 Phe Leu Leu Ala Leu Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala Tyr
4259 180 185 190
4260 Glu Val Arg Asn Val Ser Gly Ile Tyr His Val Thr Asn Asp Cys Ser
4261 195 200 205
4262 Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Val Ile Met His Thr Pro
4263 210 215 220
4264 Gly Cys Val Pro Cys Val Arg Glu Gly Asn Ser Ser Arg Cys Trp Val
4265 225 230 235 240
4266 Ala Leu Thr Pro Thr Leu Ala Ala Arg Asp Ala Ser Val Pro Thr Thr
4267 245 250 255
4268 Thr Ile Arg Arg His Val Asp Leu Leu Val Gly Thr Ala Ala Phe Cys
4269 260 265 270
4270 Ser Ala Met Tyr Val Gly Asp Leu Cys Gly Ser Ile Phe Leu Val Ser
4271 275 280 285
```